DOCUMENT RESUME

ED 410 316 TM 027 825

AUTHOR	Bond, Linda A.
TITLE	Norm- and Criterion-Referenced Testing. ERIC/AE Digest.
INSTITUTION	ERIC Clearinghouse on Assessment and Evaluation, Washington, DC.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-TM-96-09
PUB DATE	Dec 96
NOTE	4p.
CONTRACT	RR93002002
AVAILABLE FROM	ERIC Clearinghouse on Assessment and Evaluation, 210 O'Boyle Hall, The Catholic University of America, Washington, DC 20064; toll free telephone: 800-464-3742.
PUB TYPE	ERIC Publications (071)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Achievement Tests; *Criterion Referenced Tests; Elementary Secondary Education; National Norms; *Norm Referenced Tests; Selection; *Standardized Tests; *State Programs; Test Content; Test Norms; *Test Use; Testing Programs
IDENTIFIERS	ERIC Digests

ABSTRACT

Tests can be categorized into two major groups: norm-referenced tests (NRTs) and criterion-referenced tests (CRTs). NRTs are designed to highlight achievement differences between and among students to produce a dependable rank order of students across a continuum of achievement from high achievers to low achievers. With NRTs, a representative group of students is given a test before it is made available to the public. The scores of students who take the test after publication are then compared to those of the norm group. CRTs determine what test takers know and can do, not how they compare to others. They report how well students are doing relative to a predetermined performance level on a specified set of educational goals and outcomes. Both NRTs and CRTs can be standardized. Test content is an important factor in choosing between NRTs and CRTs. The content of the NRT is selected according to how well it ranks students, while the content of a CRT is determined by how well it matches the learning outcomes considered most important. It is easier to ensure the match to expected skills with a CRT. In 1994, 31 states administered NRTs and 33 states administered CRTs. Of these states, 22 administered both types of test. Before a state can choose the type of standardized test to use, state education officials will have to consider three criteria: (1) whether the assessment strategies of a test match the state's goals; (2) the content the state wishes to assess; and (3) the kinds of interpretations state officials want to make about student performance. (Contains six references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 410 316

Norm- and Criterion-referenced Testing

Linda A. Bond

ERIC/AE Digest Series EDO-TM-96-09

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM027825

BEST COPY AVAILABLE

Norm- and Criterion-referenced Testing

Linda A. Bond
Regional Policy Information Center North Central Regional Educational Laboratory

Tests can be categorized into two major groups: norm- referenced tests and criterion-referenced tests. These two tests differ in their intended purposes, the way in which content is selected, and the scoring process which defines how the test results must be interpreted. This brief paper will describe the differences between these two types of assessments and explain the most appropriate uses of each.

Intended Purposes

The major reason for using a norm-referenced tests (NRT) is to classify students. NRTs are designed to highlight achievement differences between and among students to produce a dependable rank order of students across a continuum of achievement from high achievers to low achievers (Stiggins, 1994). School systems might want to classify students in this way so that they can be properly placed in remedial or gifted programs. These types of tests are also used to help teachers select students for different ability level reading or mathematics instructional groups.

With norm-referenced tests, a representative group of students is given the test prior to its availability to the public. The scores of the students who take the test after publication are then compared to those of the norm group. Tests such as the California Achievement Test (CTB/McGraw-Hill), the Iowa Test of Basic Skills (Riverside), and the Metropolitan Achievement Test (Psychological Corporation) are normed using a national sample of students. Because norming a test is such an elaborate and expensive process, the norms are typically used by test publishers for 7 years. All students who take the test during that seven year period have their scores compared to the original norm group.

While norm-reference tests ascertains the rank of students, criterion-referenced tests (CRTs) determine *"...what test takers can do and what they know, not how they compare to others"* (Anastasi, 1988, p. 102). CRTs report how well students are doing relative to a pre-determined performance level on a specified set of educational goals or outcomes included in the school, district, or state curriculum.

Educators or policy makers may choose to use a CRT when they wish to see how well students have learned the knowledge and skills which they are expected to have mastered. This information may be used as one piece of information to determine how well the student is learning the desired curriculum and how well the school is teaching that curriculum.

Both NRTs and CRTs can be standardized. The U.S. Congress, Office of Technology Assessment (1992) defines a standardized test as *"one that uses uniform procedures for administration and scoring in order to assure that the results from different people are comparable. Any kind of test -- from multiple choice to essays to oral examinations -- can be standardized if uniform scoring and administration are used"*(p. 165). This means that the comparison of student scores is possible. Thus, it can be assumed that two students who receive the identical scores on the same standardized test demonstrate corresponding levels of performance. Most national, state and district tests are standardized so that every score can be interpreted in a uniform manner for all students and schools.

Selection of Test Content

Test content is an important factor choosing between an NRT test and a CRT test. The content of an NRT test is selected according to how well it ranks students from high achievers to low. The content of a CRT test is determined by how well it matches the learning outcomes deemed most important. Although no test can measure everything of importance, the content selected for the CRT is selected on the basis of its significance in the curriculum while that of the NRT is chosen by how well it discriminates among students.

Any national, state or district test communicates to the public the skills that students should have acquired as well as the levels of student performance that are considered satisfactory. Therefore, education officials at any level should carefully consider content of the test which is selected or developed. Because of the importance placed upon high scores, the content of a standardized test can be very influential in the development of a school's curriculum and standards of excellence.

ERIC® ERIC Clearinghouse on Assessment and Evaluation, 210 O'Boyle Hall, The Catholic University of America, Washington, DC 20064 • 800 464-3742 CUA

NRTs have come under attack recently because they traditionally have proportedly focused on low level, basic skills. This emphasis is in direct contrast to the recommendations made by the latest research on teaching and learning which calls for educators to stress the acquisition of conceptual understanding as well as the application of skills. The National Council of Teachers of Mathematics (NCTM) has been particularly vocal about this concern. In an NCTM publication (1991), Romberg (1989) cited that "*a recent study of the six most commonly used commercial achievement tests found that at grade 8, on average, only 1 percent of the items were problem solving while 77 percent were computation or estimation*" (p. 8).

In order to best prepare their students for the standardized achievement tests, teachers usually devote much time to teaching the information which is found on the standardized tests. This is particularly true if the standardized tests are also used to measure an educator's teaching ability. The result of this pressure placed upon teachers for their students to perform well on these tests has resulted in an emphasis on low level skills in the classroom (Corbett & Wilson, 1991). With curriculum specialists and educational policy makers alike calling for more attention to higher level skills, these tests may be driving classroom practice in the opposite direction of educational reform.

Test Interpretation

As mentioned earlier, a student's performance on an NRT is interpreted in relation to the performance of a large group of similar students who took the test when it was first normed. For example, if a student receives a percentile rank score on the total test of 34, this means that he or she performed as well or better than 34% of the students in the norm group. This type of information can useful for deciding whether or not students need remedial assistance or is a candidate for a gifted program. However, the score gives little information about what the student actually knows or can do. The validity of the score in these decision processes depends on whether or not the content of the NRT matches the knowledge and skills expected of the students in that particular school system.

It is easier to ensure the match to expected skills with a CRT. CRTs give detailed information about how well a student has performed on each of the educational goals or outcomes included on that test. For instance, "*. . . a CRT score might describe which arithmetic operations a student can perform or the level of reading difficulty he or she can comprehend*" (U.S. Congress, OTA, 1992, p. 170). As long as the content of the test matches the content that is considered important to learn, the CRT gives the student, the teacher, and the parent more information about how much of the valued content has been learned than an NRT.

Summary

Public demands for accountability, and consequently for high standardized tests scores, are not going to disappear. In 1994, thirty-one states administered NRTs, while thirty-three states administered CRTs. Among these states, twenty-two administered both. Only two states rely on NRTs exclusively, while one state relies exclusively on a CRT. Acknowledging the recommendations for educational reform and the popularity of standardized tests, some states are designing tests that "*reflect, insofar as possible, what we believe to be appropriate educational practice*" (NCTM, 1991, p.9). In addition to this, most states also administer other forms of assessment such as a writing sample, some form of open-ended performance assessment or a portfolio (CCSSO/NCREL, 1994).

Before a state can choose what type of standardized test to use, the state education officials will have to consider if that test meets three standards. These criteria are whether the assessment strategy(ies) of a particular test matches the state's educational goals, addresses the content the state wishes to assess, and allows the kinds of interpretations state education officials wish to make about student performance. Once they have determined these three things, the task of choosing between the NRT and CRT will becomes easier.

References

Anastasi, A. (1988). Psychological Testing. New York, New York: MacMillan Publishing Company.

Corbett, H.D. & Wilson, B.L. (1991). Testing, Reform and Rebellion. Norwood, New Jersey: Ablex Publishing Company.

Romberg, T.A., Wilson, L. & Mamphono Khaketla (1991).. "The Alignment of Six Standardized Tests with NCTM Standards", an unpublished paper, University of Wisconsin-Madison. In Jean Kerr Stenmark (ed; 1991). Mathematics Assessment: Myths, Models, Good Questions, and Practical Suggestions. The National Council of Teachers of Mathematics (NCTM)

Stenmark, J.K (ed; 1991). Mathematics Assessment: Myths, Models, Good Questions, and Practical Suggestions. Edited by. Reston, Virginia: The National Council of Teachers of Mathematics (NCTM)

Stiggins, R.J. (1994). Student-Centered Classroom Assessment. New York: Merrill

U.S. Congress, Office of Technology Assessment (1992). Testing in America's Schools: Asking the Right Questions. OTA-SET-519 (Washington, D.C.: U.S. Government Printing Office)

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract RR93002002. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**